JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2022JG007259

Characterizing Performance of Freshwater Wetland Methane Models Across Time Scales at FLUXNET-CH_4 Sites Using Wavelet Analyses

Zhen Zhang[1,2], Sheel Bansal[3], Kuang-Yu Chang[4], Etienne Fluet-Chouinard[5], Kyle Delwiche[6], Mathias Goeckede[7], Adrian Gustafson[8], Sara Knox[9], Antti Leppänen[10], Licheng Liu[11], Jinxun Liu[12], Avni Malhotra[13], Tiina Markkanen[10], Gavin McNicol[14], Joe R. Melton[15], Paul A. Miller[8], Changhui Peng[16], Maarit Raivonen[10], William J. Riley[4], Oliver Sonnentag[17], Tuula Aalto[10], Rodrigo Vargas[18], Wenxin Zhang[8], Qing Zhu[4], Qiuan Zhu[19], Qianlai Zhuang[11], Lisamarie Windham-Myers[20], Robert B. Jackson[21], and Benjamin Poulter[22]

[1]National Tibetan Plateau Data Center (TPDC), State Key Laboratory of Tibetan Plateau Earth System, Environment and Resource (TPESER), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China, [2]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA, [3]Northern Prairie Wildlife Research Center, U.S. Geological Survey, Jamestown, ND, USA, [4]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [5]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland, [6]Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, CA, USA, [7]Department of Biogeochemical Signals, Max Planck Institute for Biogeochemistry, Jena, Germany, [8]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden, [9]The University of British Columbia, Vancouver, BC, Canada, [10]Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki, Finland, [11]Department of Earth, Atmospheric, and Planetary Sciences, Department of Agronomy, Purdue University, West Lafayette, IN, USA, [12]Western Geographic Science Center, U.S. Geological Survey, Moffett Field, CA, USA, [13]Department of Geography, University of Zurich, Zurich, Switzerland, [14]Department of Earth and Environmental Sciences, University of Illinois Chicago, Chicago, WI, USA, [15]Environment and Climate Change Canada, Climate Research Division, Victoria, BC, Canada, [16]Department of Biology Sciences, Institute of Environment Science, University of Quebec at Montreal, Montreal, QC, Canada, [17]Département de géographie, Université de Montréal, Montréal, QC, Canada, [18]Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA, [19]College of Hydrology and Water Resources, Hohai University, Nanjing, China, [20]Water Mission Area, U.S. Geological Survey, Menlo Park, CA, USA, [21]Department of Earth System Science, Stanford University, Stanford, CA, USA, [22]Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Abstract Process-based land surface models are important tools for estimating global wetland methane (CH_4) emissions and projecting their behavior across space and time. So far there are no performance assessments of model responses to drivers at multiple time scales. In this study, we apply wavelet analysis to identify the dominant time scales contributing to model uncertainty in the frequency domain. We evaluate seven wetland models at 23 eddy covariance tower sites. Our study first characterizes site-level patterns of freshwater wetland CH_4 fluxes (FCH_4) at different time scales. A Monte Carlo approach was developed to incorporate flux observation error to avoid misidentification of the time scales that dominate model error. Our results suggest that (a) significant model-observation disagreements are mainly at multi-day time scales (<15 days); (b) most of the models can capture the CH_4 variability at monthly and seasonal time scales (>32 days) for the boreal and Arctic tundra wetland sites but have significant bias in variability at seasonal time scales for temperate and tropical/subtropical sites; (c) model errors exhibit increasing power spectrum as time scale increases, indicating that biases at time scales <5 days could contribute to persistent systematic biases on longer time scales; and (d) differences in error pattern are related to model structure (e.g., proxy of CH_4 production). Our evaluation suggests the need to accurately replicate FCH_4 variability, especially at short time scales, in future wetland CH_4 model developments.

Plain Language Summary Land surface models are useful tools to estimate and predict wetland methane (CH_4) flux but there is no evaluation of modeled CH_4 flux error at different time scales. Here we use a statistical approach and observations from eddy covariance sites to evaluate the performance of seven wetland models for different wetland types. The results suggest models have captured CH_4 flux variability at monthly or seasonal time scales for boreal and Arctic tundra wetlands but failed to capture the observed seasonal variability for temperate and tropical/subtropical wetlands. The analysis suggests that improving modeled flux at short time scale is important for future model development.

Key Points:

- Significant model-observation disagreements were found at multi-day and weekly time scales (<15 days)
- Models captured variability at monthly and seasonal time (42–142 days) scales for boreal and Arctic tundra sites but not for temperate and tropical sites
- The model errors show that biases at multi-day time scales may contribute to persistent systematic biases on longer time scales

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Z. Zhang,
zzhang88@umd.edu

Citation:

Zhang, Z., Bansal, S., Chang, K.-Y., Fluet-Chouinard, E., Delwiche, K., Goeckede, M., et al. (2023). Characterizing performance of freshwater wetland methane models across time scales at FLUXNET-CH_4 sites using wavelet analyses. *Journal of Geophysical Research: Biogeosciences*, *128*, e2022JG007259. https://doi.org/10.1029/2022JG007259

Received 24 OCT 2022
Accepted 10 OCT 2023
Corrected 8 FEB 2024

This article was corrected on 8 FEB 2024. See the end of the full text for details.

Author Contributions:

Conceptualization: Zhen Zhang, Rodrigo Vargas, Lisamarie Windham-Myers

Data curation: Kyle Delwiche, Mathias Goeckede, Adrian Gustafson, Antti Leppänen, Licheng Liu, Tiina Markkanen, Gavin McNicol, Joe R. Melton, Paul A. Miller, Changhui Peng, Maarit Raivonen, William J. Riley, Oliver Sonnentag, Tuula Aalto, Rodrigo Vargas, Wenxin Zhang, Qing Zhu, Qiuan Zhu, Qianlai Zhuang
Formal analysis: Zhen Zhang
Funding acquisition: Lisamarie Windham-Myers, Robert B. Jackson, Benjamin Poulter
Investigation: Zhen Zhang, Sheel Bansal, Kuang-Yu Chang, Etienne Fluet-Chouinard, Sara Knox, Wenxin Zhang
Methodology: Zhen Zhang
Project Administration: Robert B. Jackson, Benjamin Poulter
Resources: Kyle Delwiche, Sara Knox, Gavin McNicol, Joe R. Melton, Changhui Peng, William J. Riley, Tuula Aalto, Qianlai Zhuang, Lisamarie Windham-Myers, Robert B. Jackson
Software: Zhen Zhang
Validation: Zhen Zhang
Visualization: Zhen Zhang
Writing – original draft: Zhen Zhang, Sheel Bansal, Kuang-Yu Chang, Etienne Fluet-Chouinard, Kyle Delwiche, Mathias Goeckede, Adrian Gustafson, Sara Knox, Tiina Markkanen, Gavin McNicol, Joe R. Melton, Paul A. Miller, Changhui Peng, Maarit Raivonen, William J. Riley, Oliver Sonnentag, Tuula Aalto, Rodrigo Vargas, Wenxin Zhang, Qing Zhu, Qiuan Zhu, Qianlai Zhuang, Lisamarie Windham-Myers, Benjamin Poulter

1. Introduction

Methane (CH_4) is a potent important greenhouse gas in terms of radiative forcing whose concentration in the atmosphere (~1,900 ppb) has increased by approximately 150% since pre-industrial times (Canadell et al., 2021; IPCC, 2013). Methane emitted from wetland ecosystems is the largest natural source at ~120–180 Tg CH_4 yr^{-1} (Poulter et al., 2017; Saunois et al., 2020) and contributes to short-term trend and interannual variability observed in atmospheric CH_4 concentration (Bousquet et al., 2006; Saunois et al., 2017; Zhang, Fluet-Chouinard, et al., 2021; Zhang, Poulter, et al., 2021). Our understanding of global wetland CH_4 emissions heavily relies on process-based wetland CH_4 models, which incorporate biogeochemical mechanisms, climate forcing variables (e.g., temperature), and spatio-temporal distributions of surface inundation and wetland extent across the world (Melton et al., 2013; Wania et al., 2013; Xu, Riley, et al., 2016; Xu, Yuan, et al., 2016; Zhang, Fluet-Chouinard, et al., 2021; Zhang, Poulter, et al., 2021). These models play a critical role in diagnosing and forecasting terrestrial CH_4 dynamics across space and time, but their wetland CH_4 flux (FCH_4) estimates have large uncertainties due to potential biases in parameterizations, limited mechanistic characterization of known CH_4 processes such as thermal impact of rainfall (Neumann et al., 2019) and microbial dynamics on FCH_4 (Chadburn et al., 2020). However, it is unclear how well the current wetland models can replicate the observed FCH_4 variability and magnitude at different time scales. Therefore, it is necessary to evaluate wetland CH_4 model performance against observations to identify temporal model error patterns and inform future model development.

So far there has not been a major synthesis effort to evaluate multiple wetland CH_4 models against global coverage of eddy covariance (EC) observations from different biomes using a standard simulation protocol, despite a few efforts to evaluate a single wetland CH_4 model at multiple sites (Ringeval et al., 2014; Wania et al., 2010) and a model inter-comparison (Melton et al., 2013; Wania et al., 2013). Moreover, current models may have a biased seasonal cycle over temperate and tropical wetlands, as suggested by a few recent regional studies (Lunt et al., 2019; Maasakkers et al., 2021; Yu et al., 2021). It is necessary to evaluate the simulated FCH_4 variability against EC observations for different biomes at different time scales. Arguably, model development to represent terrestrial CH_4 processes has been hindered by (a) limited number of local-to-regional CH_4 observations to evaluate model outputs; (b) lack of understanding of the underlying processes informed by EC measurements and how well these processes are represented in the models. Evaluations of wetland CH_4 models against the recently complied database FLUXNET-CH_4 (Delwiche et al., 2021; Knox et al., 2019, 2021) offer an opportunity to improve understanding of current model performance for different wetland types.

Despite previous observational synthesis studies (Chang et al., 2021; Delwiche et al., 2021; Knox et al., 2021) that have identified the major controlling factors that regulate temporal variations in freshwater wetland FCH_4 at different time scales, it is currently unknown how accurate wetland CH_4 models are in predicting FCH_4 across time scales and what factors are likely causing model biases across different temporal scales. Knox et al. (2021) demonstrated that dominant factors controlling the seasonality in EC-based FCH_4 vary with wetland types and the major processes that regulate FCH_4 vary at different time scales (e.g., from sub-daily to seasonal). For example, although soil temperature simulations are well established in wetland models with different thermal parameterization schemes, the representation of the modeled relationship between FCH_4 and temperature should be closely evaluated since it may affect model performance for the high latitudes. Examples include cold regions influenced by freeze-thaw cycles where FCH_4 may occur during the zero-curtain period when subsurface soil temperature are poised near 0°C (Tao et al., 2021; Zona et al., 2016). In addition, temperature hysteresis could contribute to different FCH_4 drivers across seasons (Chang et al., 2021). In contrast, models tend to use different proxies to calculate microbial CH_4 production (e.g., Gross Primary Production, Net Primary Production [NPP], ecosystem Heterotrophic Respiration [Reco], and carbon substrate concentrations), which likely influences simulated accuracy in reproducing FCH_4 at different time scales.

It is difficult to diagnose the mechanisms responsible for the lack of agreement between model and observation using conventional model-fitting approaches (Schaefer et al., 2012; Taylor, 2001) that apply statistical metrics (e.g., RMSE, r^2, standard deviation). In contrast, model-observation evaluations in the frequency domain using wavelet analysis (Figure 1) or Fourier transform can provide insights about model-observation disagreements at different temporal scales (Dietze et al., 2011; Stoy et al., 2013; Vargas et al., 2010). Wavelet analysis is especially useful for model evaluation since, compared to Fourier transform, it can identify not only the time scales that influence a signal but also inform when those time scales are significant. Previous studies have identified disagreement between models and observations for carbon dioxide (CO_2) fluxes across different ecosystems

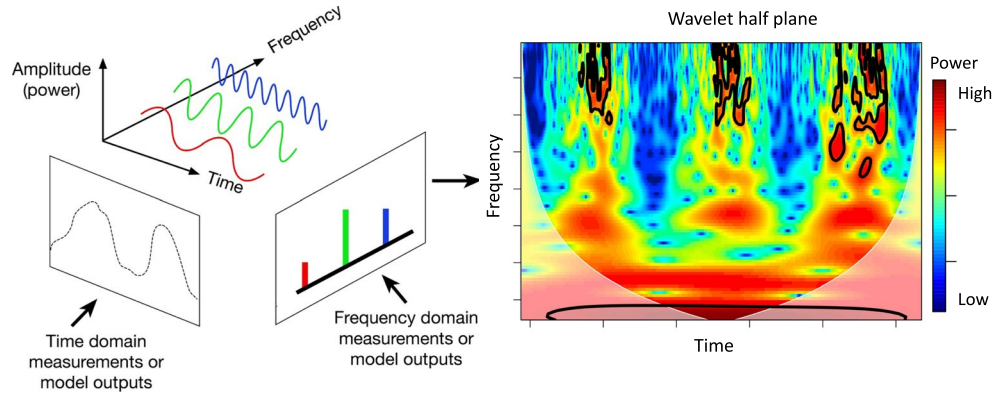

Figure 1. A conceptual description of differences between information in the time-domain and the frequency domain and an example of model-data evaluation in the frequency domain (adapted from Vargas et al. (2010)). A time series can be decomposed into time and frequency (i.e., time scale) domain using the continuous wavelet transform. The resulting wavelet power spectra are plotted on what is referred to as wavelet half-plane, where time is along *x* axis, frequency/time scale along *y* axis, and spectra power indicated by color. Statistically significant areas are clearly distinguished by thick black contour lines. The white line is the cone of influence (COI) beyond which wavelet coefficients are unreliable (referred to as "edge effect"). While the COI is included in the calculations, we do not draw any conclusions or interpretations based on the wavelet coefficients beyond the COI.

(Golub et al., 2023; Li et al., 2023; Richardson et al., 2012; Schwalm et al., 2010; Stoy et al., 2013) using wavelet analysis, and have found that (a) model errors in CO_2 flux peak at the diurnal scale and (b) the model error patterns are associated with model structure and environmental drivers. However, so far there is no assessment for FCH_4. Consequently, assessments of model-observation agreements using wavelet analysis are needed to identify discrepancies between observed and modeled FCH_4 and provide insights for model development.

Our study aims to evaluate the performance of wetland CH_4 models in the frequency domain against a large ensemble of EC measurements of ecosystem-scale FCH_4. The goal is to quantify the most important time scales (e.g., multi-day, monthly, and seasonal) for the variability of FCH_4 across wetland types and provide insights about the time periods in which models should be improved. Our specific objectives are to: (a) quantify the most relevant time-scales for the variability of FCH_4 in the models and observations at the site-level, (b) test the disagreement between in situ observations and modeled FCH_4 in time-frequency domain, (c) give insights into model structures responsible for model/observation mismatch. Based on previous findings for CO_2 flux (Dietze et al., 2011; Stoy et al., 2013), we hypothesize that (a) models will have better model-observation agreement in terms of flux variability at longer time scales (e.g., monthly to seasonal) than short to intermediate time scale (e.g., multi-day to sub-monthly) as important biological processes regulated by seasonal variation (e.g., CH_4 production response to temperature) are adequately formulated in the models; (b) models will tend to fail at multi-day and sub-monthly time scales due to forcing error propagation and limited representation of modeled plant physiology and biogeochemical processes; (c) The models have better performance over boreal and Arctic tundra sites than temperate and tropical sites, as temperature become less dominating in controlling FCH_4 variability for those wetland types.

2. Materials and Methods

We used data from 23 freshwater wetland sites included in the FLUXNET-CH_4 Community Product (Delwiche et al., 2021) to evaluate seven wetland CH_4 models from the Global Carbon Project (GCP) Methane Budget (Saunois et al., 2020; Stavert et al., 2021). The model simulations follow a common simulation protocol using a gridded climate data set from Climate Research Unit (CRU/CRU-JRA; CRU-JRA is a 6-hourly interpolated climate data set from Japanese Reanalysis data; JRA, i.e., aligned with CRU on the monthly basis) as the inputs. The selected EC sites have a total of 70 site-years of data classified as boreal forest/taiga (n = 25), Arctic tundra (n = 15), temperate (n = 25), and tropical/subtropical wetlands (n = 5). We take into account the flux measurement errors in identifying model-data disagreements with observations by assessing the contribution of flux-tower observations error via a Monte Carlo approach.

Table 1
Summary of Site Characteristics

Site ID	Country	Latitude	Longitude	Biome type	Wetland type	Start year	End year	Data reference
BR-NPW	Brazil	−16.50	−56.41	Tropical/Subtropical	Seasonal	2014	2016	Dalmagro et al. (2019)
CA-SCB	Canada	61.31	−121.30	Boreal forest	Bog	2014	2018	Sonnentag and Helbig (2020)
FI-LOM	Finland	68.00	24.21	Boreal forest	Fen	2008	2010	Lohila et al. (2020)
MY-MLM	Malaysia	1.46	111.15	Tropical/Subtropical	Swamp	2014	2015	Wong et al. (2020)
RU-VRK	Russia	67.06	62.94	Arctic tundra	Wet Tundra	2008	2008	Friborg and Shurpali (2020)
SE-DEG	Sweden	64.18	19.56	Boreal forest	Fen	2014	2017	Nilsson and Peichl (2020)
SE-ST1	Sweden	68.35	19.05	Arctic tundra	Fen	2012	2014	Jansen et al. (2020)
SE-STO	Sweden	68.36	19.05	Arctic tundra	Bog	2015	2015	Jansen et al. (2020)
US-ATQ	USA	70.47	−157.41	Arctic tundra	Wet Tundra	2014	2014	Zona and Oechel (2020a)
US-BZB	USA	64.70	−148.32	Boreal forest	Bog	2014	2016	Euskirchen and Edgar (2020)
US-BZF	USA	64.70	−148.31	Boreal forest	Fen	2014	2016	Euskirchen (2022a)
US-BZS	USA	64.70	148.32	Boreal forest	Swamp	2015	2016	Euskirchen (2022b)
US-EML	USA	68.88	−149.25	Arctic tundra	Bog	2015	2016	Schuur (2020)
US-ICS	USA	68.61	−149.31	Arctic tundra	Wet Tundra	2015	2016	Euskirchen et al. (2020)
US-IVO	USA	68.49	−155.75	Arctic tundra	Wet Tundra	2013	2016	Zona and Oechel (2020b)
US-LOS	USA	46.08	−89.98	Temperate	Fen	2014	2017	Desai and Thom (2020)
US-ORV	USA	40.02	−83.02	Temperate	Marsh	2011	2015	Bohrer and Morin (2020)
US-OWC	USA	41.38	−82.51	Temperate	Marsh	2015	2015	Bohrer and Morin (2020)
US-SNE	USA	38.04	−121.75	Temperate	Marsh	2016	2017	Shortt et al. (2020)
US-TW1	USA	38.11	−121.65	Temperate	Marsh	2011	2017	Valach et al. (2020)
US-UAF	USA	64.87	−147.86	Boreal forest	Bog	2011	2017	Iwata et al. (2020)
US-WPT	USA	41.47	−82.99	Temperate	Marsh	2011	2013	Chen and Chu (2020)

2.1. FLUXNET-CH_4

Twenty-three sites from the FLUXNET-CH_4 database were selected for the analysis (Table 1; Figure 2) based on three criteria: (a) tidal, upland, and agricultural sites were excluded from the analysis as the models only simulate natural inland freshwater wetland FCH_4; (b) there had to be at least one complete site-year of overlapping EC observations and model results for all seven wetland models; and (c) only restored freshwater wetlands at later stages of wetland development (>10 years) were included in the analysis.

In order to match the broad definition of freshwater wetlands in the models, selected EC sites were regrouped to represent a broad range of wetland/biome type along a latitudinal gradient. The original freshwater wetland types were classified into bog, fen, marsh, and swamp based on site-specific literature (Delwiche et al., 2021). The biome types (Arctic tundra, boreal forest/taiga, temperate, and tropical/subtropical), were defined based upon Olson et al. (2001) using site coordinates and vegetation types to group wetland sites. Since continuous wavelet decomposition requires a gap-free time series, we used gap-filled data from the FLUXNET-CH_4 database. Details on data standardization and gap-filling are described in Knox et al. (2019) and Delwiche et al. (2021). Gaps in

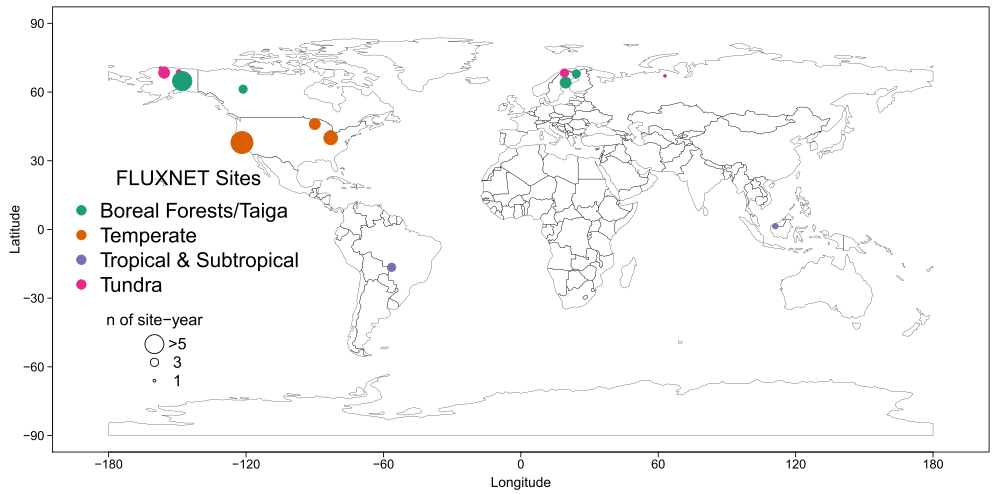

Figure 2. Locations of eddy covariance sites from FLUXNET-CH_4 in this study, with sites colored by wetland types. The variable size of dots in the map corresponds to the sample size (i.e., number of site-years) with a larger dot indicating a higher sample size. Base map used came from https://hub.arcgis.com/datasets/esri::world-countries.

FCH_4 were filled using artificial neural networks (ANN) as described in Knox et al. (2019). An estimate of FCH_4 observation error at every time step was generated, accounting for uncertainties associated with the gap-filling process and random measurement errors. The errors of observed FCH_4 follows a double exponential distribution (Knox et al., 2019), which has a fatter tail than normal and is highly heteroscedastic, with error increasing linearly with the absolute magnitude of the flux, similar to CO_2 flux errors as suggested by previous studies (Hollinger & Richardson, 2005; Lasslop et al., 2008; Richardson et al., 2006, 2008). These uncertainties are incorporated in the spectral null model, as described in the spectral analysis (Section 2.3).

2.2. Wetland FCH_4 Models

Our study applies seven global wetland CH_4 models from the GCP Methane Budget activities (Saunois et al., 2020), including Canadian Land Surface Scheme Including Biogeochemical Cycles (CLASSIC), E3SM Land Model (ELM), Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS), Jena Scheme for Biosphere-Atmosphere Coupling in Hamburg HelsinkI Model of MEthane buiLd-up and emission (JSBACH-HIMMELI), Lund-Potsdam-Jena *Wald Schnee and Landschaft* version (LPJ-wsl), Terrestrial Ecosystem Model-Methane Dynamics Module, and TRIPLEX-GHG. The details about the structure and configurations of the wetland CH_4 models can be found in Table 2. All the models were run to steady-state using their own parameters and no site-specific tunings were done. Ancillary data such as soil texture and CH_4-related parameter sets were used as model-specific inputs (Table 2). Thus the assumptions about the local environment at each site depended on the individual model's setup. The models were run at the global scale at their native spatial resolution following a prescribed protocol to facilitate intercomparison. The models were run at the grid cell level using the CRU-JRA 6-hourly, land surface, gridded climate data set, which was constructed by combining the CRU data set and the reanalysis from Japanese Reanalysis data (JRA) produced by the Japanese Meteorological Agency. The CRU-JRA was adjusted where possible to align with the monthly climate data set CRU (version ts3.26) data. For CLASSIC that requires climate inputs at half-hour time step, the CRU-JRA has been interpolated to half-hourly using a linear interpolation or random distributions. LPJ-wsl model uses the monthly CRU data set, and a weather generator within the model to produce precipitation events and daily temperature. Here we evaluate the wetland FCH_4 strength ($gCH_4\ m^{-2}\ day^{-1}$), which was defined as the total flux over a 24-hr period over a standardized wetland area (m^2). The FCH_4 strength is calculated as FCH_4 divided by wetland areal fraction within the grid cell to exclude the effect of inundation dynamics in the FCH_4 calculation.

The wetland CH_4 models can be generally characterized as a set of functions describing the biogeochemical processes that control CH_4 production and oxidation through methanogenesis and methanotrophy, and the biophysical processes that regulate CH_4 transport from the soil to the atmosphere (Table 2). Methanogenesis in the models is linked to different proxies (e.g., carbon substrate, heterotrophic respiration, NPP) with a wide range

Journal of Geophysical Research: Biogeosciences 10.1029/2022JG007259

Table 2
Summary of Model Characteristics

Model	Wetland PFT	Components of CH_4 flux	Temperature response functions	CH_4 production proxy	Nitrogen cycles	Fire	Spatial resolution	Forcing time step	Reference
CLASSIC	No wetland-specific PFTs	Net flux	Indirectly through Rh (see Section A3.2 in Melton and Arora (2016))	Rh is scaled to account for CH_4 versus CO_2 emitted and differences in upland versus lowland Rh	No	Yes	T63 (~2.8)	30 min	Arora et al. (2018) and Melton and Arora (2016)
ELM	No wetland-specific PFTs	Gross production; gross consumption; oxidation; diffusive, aerenchyma, and ebullition fluxes	Q10 based on soil T in each soil layer	Rh in each soil layer is scaled to estimate CH_4 production	Yes	Yes	~2°	6-Hourly	Riley et al. (2011), Xu, Yuan, et al. (2016), and Xu, Riley, et al. (2016)
JSBACH-HIMMELI	Generic wetland PFT with C3 grass parameters for vegetation	Net wetland soil CH_4 flux	Layered soil temperature, different temperature responses for production, consumption, diffusion	CH_4 production depends on anoxic respiration produced by YASSO soil carbon model modified to account for anoxic conditions and coupled to JSBACH	No	No	1.875°	Daily	Raivonen et al. (2017)
LPJ-wsl	No wetland-specific PFTs	Net flux	Soil temperature calculation in LPJ is 12 layers scheme following Wania et al. (2009). Daily average soil temperature for 0–50 cm depth is used for CH_4 function	Rh	No	Yes	0.5	Monthly	Zhang et al. (2016, 2018)

Journal of Geophysical Research: Biogeosciences
10.1029/2022JG007259

Table 2
Continued

Model	Wetland PFT	Components of CH_4 flux	Temperature response functions	CH_4 production proxy	Nitrogen cycles	Fire	Spatial resolution	Forcing time step	Reference
LPJ-GUESS	High-latitude (>40°N): Wetland grass, cushion forbs, lichens, sphagnum moss. South of 40°N: C3 and C4 grasses only on wetlands	Net and gross emissions are simulated for high-latitude (>40°N) ecosystems, where CH_4 is released through diffusion, plant-mediated, and ebullition pathways. South of 40°N: net emissions only based on a simple rescaling of heterotrophic respiration	Decomposition of litter and SOM uses an empirical relationship for temperature response of soil temperature at 25 cm depth (calculated following Wania et al. (2009)) across ecosystems, incorporating damping of Q10 response due to temperature Acclimation. CH_4 production, oxidation and transport use temperature dependencies from Wania et al. (2010), from each 10 cm layer in the soil	CH_4 production depends on soil temperature in each 10 cm soil layer, the degree of anoxia and the availability of substrate that consists of a fraction of litter and soil carbon decomposition	Yes	No, not on wetlands	0.5	Monthly, interpolated to quasi-daily values	McGuire et al. (2012) and Wania et al. (2009, 2010)
TEM-MDM	Five primary types of wetlands are considered in boreal, temperate and tropical regions (total 15 subtypes). They are forested bog, nonforested bog, forested swamp, nonforested swamp and alluvial formations	Gross production; gross consumption; net flux	Q10 coefficient is used to account for soil temperature effects on methanotrophy rates within each 1 cm layer of the soil profile	CH_4 production is modeled as an anaerobic process that occurs in the saturated zone of the soil profile, controlled by methanogenic substrate availability, soil temperatures, PH, and redox potential	Yes	No	0.5	Daily	L. Liu et al. (2020), Zhuang et al. (2004), and Zhuang et al. (2013)
TRIPLEX-GHG	A general wetland PFT was added without considering specific wetland plants type	Net flux	Soil temperature factor was evaluated with and exponential function that considering soil temperature and optimum soil temperature for CH_4 production. The Q10 in the temperature function for CH_4 production and CH_4 oxidation could be calibrated separately	CH_4 production was calculated as a proportion of heterotrophic respiration (CO_2–C) along with soil temperature, Eh and pH modification factors	Yes	No	0.5	Daily	Zhu et al. (2015)

of model complexity—more sophisticated models include wetland Plant Functional Types (PFTs) and explicitly simulate the processes of CH_4 production, consumption, and transport, while the simplified models use generalized empirical equations to simulate net flux without considering individual components of FCH_4. More complex model structures provide capacity to capture the important temporal patterns of FCH_4 but this invariably leads to additional parameter uncertainty due to the scarcity of observational constraints. The response function of FCH_4 dynamics to temperature in each model is another factor that influences the simulated time series of FCH_4. For example, for high-latitude wetlands, model representations of freeze-thaw cycles influence the performance in capturing FCH_4 during the early spring thaw and zero-curtain period (Zona et al., 2016). It is worth noting that assessing the overall complexity of the wetland models is challenging due to its integrations with multiple processes such as freeze/thaw cycle, soil thermal schemes, nutrient cycles, and other components within land surface models. While certain wetland methane modules may appear simpler than others in terms of represented processes and parameters, it is not straightforward to establish a clear ranking of complexity.

2.3. Evaluation Strategy and Wavelet Analyses

This analysis focused on the comparison of observed and modeled FCH_4. All analyses were conducted using daily time series. Since the modeled FCH_4 fluxes are not directly comparable to the EC measurements due to the spatial mismatch between modeled gridded fluxes and site-level observations, we evaluate simulated FCH_4 by calculating the normalized residual error (NRE, $\varepsilon_{s,m,t}$) between normalized model and observation as:

$$\varepsilon_{s,m,t} = \left(\frac{\mathrm{Model}_{s,m,t} - \overline{\mathrm{Model}_{s,m,t}}}{\sigma_{s,m}} \right) - \left(\frac{\mathrm{Data}_{s,t} - \overline{\mathrm{Data}_{s,t}}}{\sigma_{s}} \right) \quad (1)$$

Where the subscripts denote site (s), model (m), and time (t) and the overbar denotes the average over the full length of the time series. The model and observation results were mean-centered to eliminate biases in the net flux, and divided by the standard deviation (σ) across the entire record to normalize the amplitude of variability. This NRE metric can be used to compare the synchrony of the model with the observation rather than evaluating absolute model biases.

We applied wavelet analysis to decompose the FCH_4 time series into an additive series of wave functions that have time scales of variability from 2 to 124 days. Wavelet analysis can identify the time scales that dominate a signal because wave functions that best match the fluctuations in the data will explain the most variance (i.e., power). Specifically, we used the continuous wavelet transform because of its ability to translate a time series into the frequency domain and its suitability for visual interpretation. The ability to discern small intervals of scales (i.e., spectral resolution) depends on the choice of the mother wavelet function (Cornish et al., 2006). For this, we applied the Morlet wavelet, a complex non-orthogonal wavelet that has been widely used for geophysical applications (Torrence & Compo, 1998) and biometeorological measurements (Meyers et al., 1993). Following a similar definition from Knox et al. (2021), time scales of variation were classified into four bands, "multi-day scale" (2–5 days), "weekly scale" (5–15 days), "monthly scale" (15–42 days), and the "seasonal scale" (>42 days). It is important to note that the "seasonal" time scale defined in our study, with an upper bound of 124 days, is notably shorter compared to the "seasonal" time scale defined in Sturtevant et al. (2016) and Knox et al. (2021). Consequently, the seasonal time scale in our study is more in line with a time scale of approximately up to 3 months. The four bands were then summarized on both a by-site and by-model basis regarding the relative contribution of each band to the overall spectra. The continuous wavelet decomposition was computed using the Morlet wavelet basis function (function name: wt) from the R package "biwavelet" (Gouhier et al., 2021). We use the bias-corrected wavelet power following Y. Liu et al. (2007) to ensure a consistent definition of power in order to enable comparisons across spectral peaks. Wavelet power spectra on very long timescales (>64 days) often exceed the so-called cone-of-influence (COI) beyond which edge effects become important due to incomplete time locality across frequencies. Therefore, the power spectra outside of COI is not interpreted here.

An appropriate null model is important to determine whether the model-observation disagreement is statistically significant. We applied a similar approach to that of Dietze et al. (2011) to generate 1,000 sets of "pseudo" time series for each site using a Monte Carlo approach. The NRE between the pseudo time series and the original data and the wavelet spectra of the NRE were calculated in the same way as the model errors. The 1,000 replicates of pseudo time series were generated with the uncertainties estimated by Knox et al. (2019) accounting for both

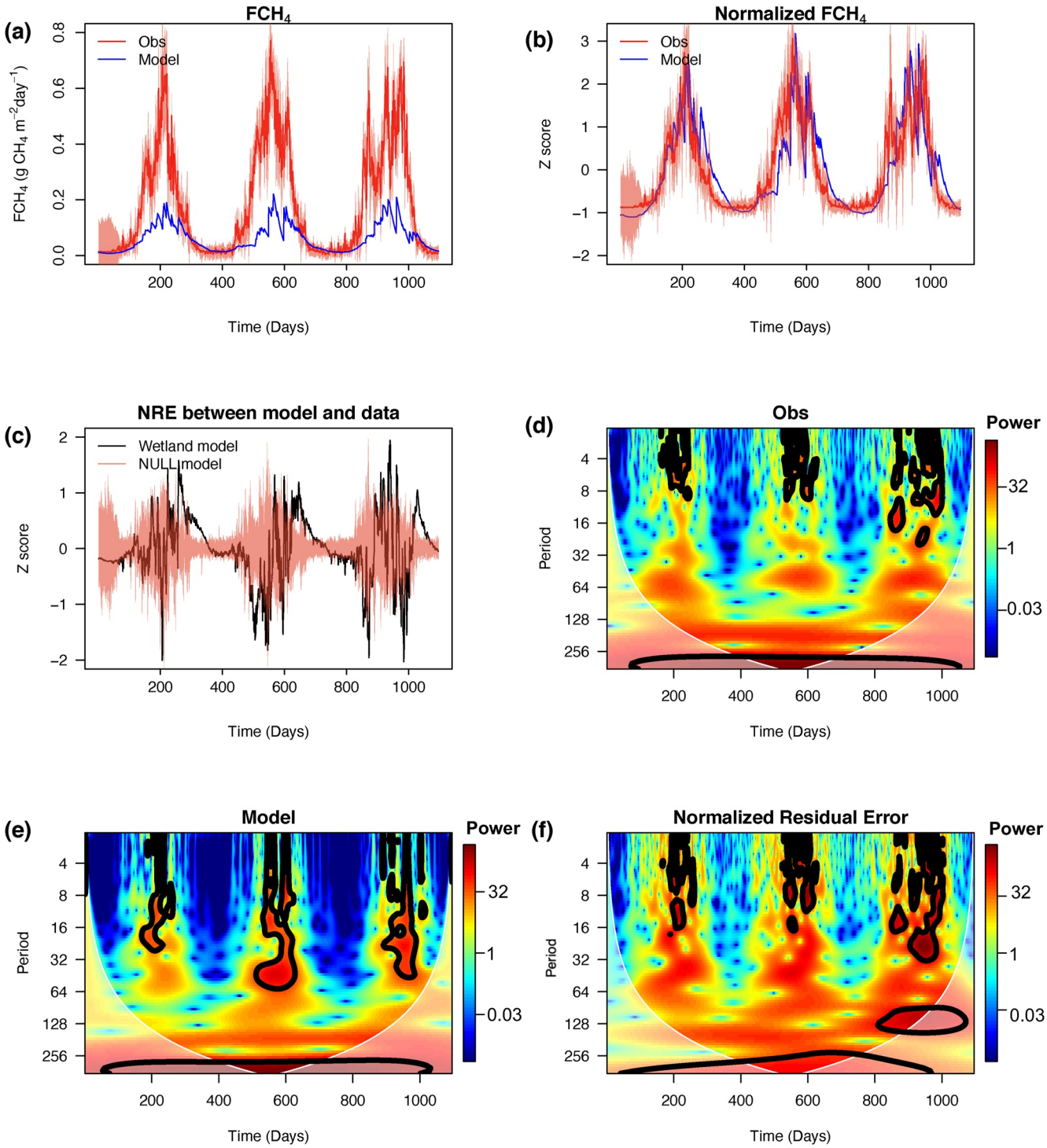

Figure 3. Example of wavelet decomposition and identification of the LPJ-wsl model error with eddy covariance observations at the US-WPT site for 2011–2013. (a) Time series of observations (Obs) methane flux (FCH_4, red line) with $1-\sigma$ observational uncertainty (shaded red area) and LPJ-wsl modeled FCH_4 (Model, blue line). (b) Normalized time series of FCH_4 from model and observations; the shaded area in red represents the upper and lower range of the normalized pseudo time series from the Monte Carlo simulations. (c) Time series of normalized residual error (NRE) (Z-score of NRE) between wetland model and observations, with shaded area in red representing NRE between observations and normalized pseudo time series, that is, NULL model. (d) Wavelet coefficients displayed in the wavelet half-plane for the normalized observations, (e) same as (d) but for LPJ-wsl modeled FCH_4 (Model), (f) NRE between model and the observations.

uncertainties in the ANN-based gap-filling algorithm and measurement uncertainty. Systematic errors due to representativeness (Chu et al., 2021; Pallandt et al., 2021), lack of nocturnal mixing, sub-mesoscale circulations, and other factors are not discussed here (Baldocchi, 2014; Peltola et al., 2015). Also note that, because the uncertainty from ANN estimation was strongly linked to the sample size, the flux errors tended to be high when the measurement availability was limited by local meteorological conditions such as the snow presence and soil freeze and thaw cycles.

The wavelet spectra were evaluated in the following ways:

- Significant spectra regions. The significant region was defined by counting the total number of area in the time-frequency distribution where the spectral characteristics of FCH_4 and model-observation mismatch were statistically significant. It was calculated by re-coding significant power as 1 and non-significant power as 0 and then stacking all site-years to count the total number.
- Marginal distribution of power spectrum of the NRE. The disagreements in the marginal power spectra were aggregated by the four defined time bands to summarize model performance across different time scales.
- Scaling exponent a for each model was calculated to quantify the spectral properties of persistence of auto-correlation structure (i.e., memory effect) in model error. This is particularly relevant because land surface models employ numerical discretization methods and numerical approximations (e.g., soil thermal schemes that affects FCH_4 calculation), which can introduce errors that accumulate and impact subsequent model predictions. Scaling exponent a was expressed as the slope of the log-log transformed relationship between frequency (i.e., time scale) and power. The scaling exponent a with a range between 1 and 2 was considered as intermediate "pink" noise between "white" and "red" noise. White or red noise indicated that if the modeled FCH_4 had a persistent memory effect (i.e., autocorrelation structure), which can be attributed to model error which resulted in larger and long-lived systematic biases at longer time scales.

One-way analysis of variance (ANOVA) was used to diagnose the relationship between model structure and the marginal distribution of spectra power for wetland types. The marginal distribution of spectral power of each band was compared with different groups of models for each wetland type. The model structures are defined in Table S1 in Supporting Information S1 to identify if there were significant differences ($p < 0.001$) between model groups.

3. Results

3.1. Wavelet Decomposition of FCH_4 Time Series From LPJ-wsl at an Example Site

Figure 3 shows the time series of FCH_4 from the observations and one model (LPJ-wsl) and demonstrates its wavelet-based power spectra at one marsh site (US-WPT) in the central U.S. (Chu et al., 2015). We use this example to explain the Monte Carlo analysis with pseudo-data and discuss the model-observation disagreement. Figure 3a shows that FCH_4 simulated by the LPJ-wsl model generally captured the seasonal cycle, but with a lower magnitude at the freshwater marsh site. The model also captured a dip in FCH_4 after the peak during the June-July-August (JJA) months, which is consistent with the observed temporal pattern. Figure 3b suggests that the temporal patterns of normalized FCH_4 between the model and observations have a good agreement ($r = 0.83$, $p < 0.05$). The relatively high uncertainty in the observed FCH_4 at the beginning of 2011 is mainly due to the limited number of observations, which causes higher uncertainty in the gap-filling method. This example shows that the discrepancies between the modeled and observed FCH_4, and the NRE uncertainty range from the null model, tend to be higher during the JJA months when the flux intensity is relatively high and highly variable, or when the data availability is limited (Figure 3c).

Both the observation and the model show significant power spectra during the JJA months (Figures 3d and 3e). The modeled FCH_4 have a longer range of dominant time scale from 2 to 64 days than the observed 2–8 days. The modeled FCH_4 has weaker spectral powers (colors toward blue) during the winter and spring seasons, indicating that the model may have less variability than the observations during the winter and spring seasons (Figures 3d and 3e). It is important to note that the power spectra of the NRE are not the difference between the wavelet coefficients displayed in Figures 3d and 3e.

The wavelet plot for the NRE suggests the largest discrepancies is mostly from JJA months, reflected as strong spectral power in the wavelet NRE (Figure 3f). It is encouraging that there is a degree of correspondence between

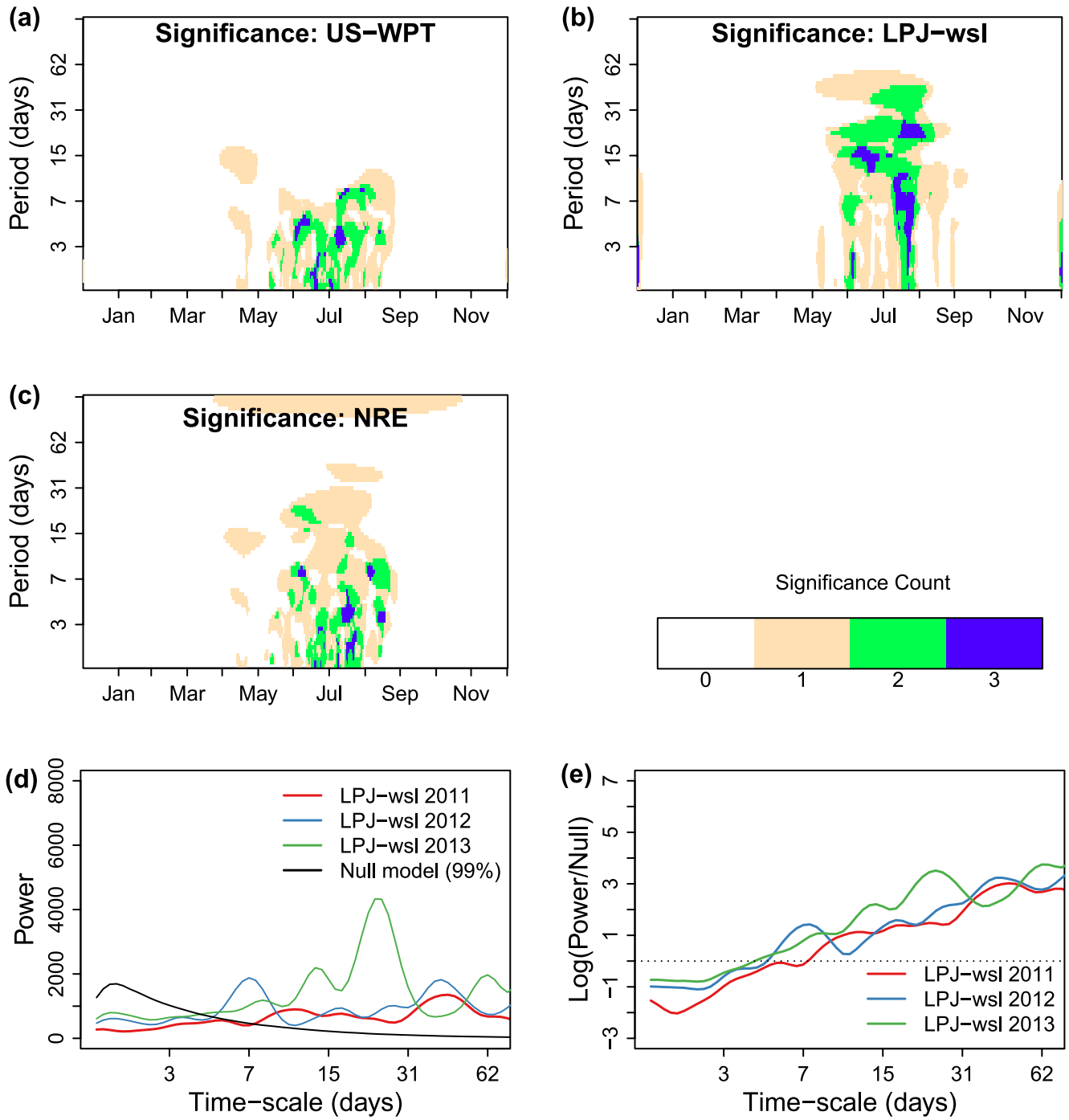

Figure 4. Wavelet evaluation of FCH_4 for the LPJ-wsl model against 3 site-year observations (2011–2013) at US-WPT site. (a) Count of significant power in the time-frequency domain for the time series of FCH_4 observations. (b) Same as (a) but for LPJ-wsl modeled FCH_4. (c) Count of significant power of normalized residual errors (NRE) between model and observations. (d) Marginal distribution of power spectra of NRE as compared to the null spectra (99% of confidence interval, solid black line) based on measurement uncertainties for each year 2011, 2012, and 2013 (red, blue, and green lines, respectively). (e) The marginal distribution of power spectra of NRE divided by the maximum of the null spectra (NULL) on a log scale. Values greater than 0 (dotted line) indicate that the model error has significantly more spectral power at those time scales than would be expected based on observation error.

the model and observations: (a) the mismatch between model and observations is not significant at the monthly and seasonal time scale (>32 days) except for 2013 when anomalously high FCH_4 is observed in August; (b) the wavelet coefficients in NRE have a low magnitude during the December-January-February (DJF) months, suggesting a less important role of the winter season fluxes at US-WPT. It is also worth noting that the seasonal

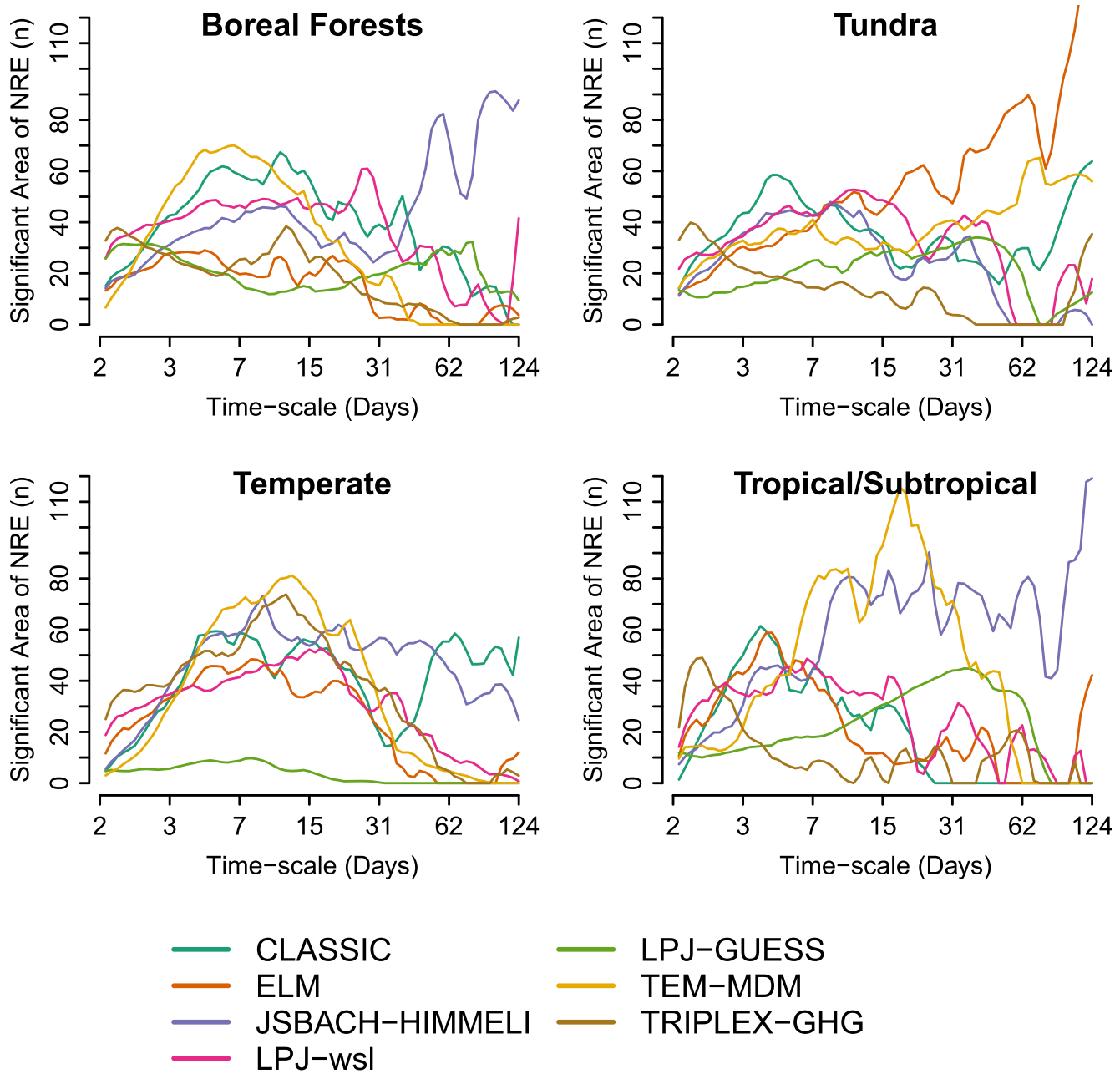

Figure 5. Significant model-observation disagreement along with time scales for all sites by biomes, represented by the marginal distribution of significant regions of normalized residual error. High values of the significant region indicate high tendency of model-observation mismatch and vice versa. The significant region is defined as the areas where the wavelet power spectrum is statistically significant (95% confidence interval). The marginal distribution of significant regions is then calculated by stacking all site-years to count the significant power in the time-frequency domain.

cycle of observed FCH_4 has much higher year-to-year variations than the modeled fluxes, which is partly due to the strong influence of local environmental conditions on the measured seasonal cycle that are not captured by the model.

3.2. Evaluation of LPJ-wsl at the Example Site

Model-observation discrepancies in LPJ-wsl at the US-WPT site were highest at daily to weekly scales. Figure 4 show that the measurements identify significant regions at high frequency (i.e., multi-day to weekly scales) while LPJ-wsl displays significant regions in the whole range of frequencies with more areas at the mid-to-low

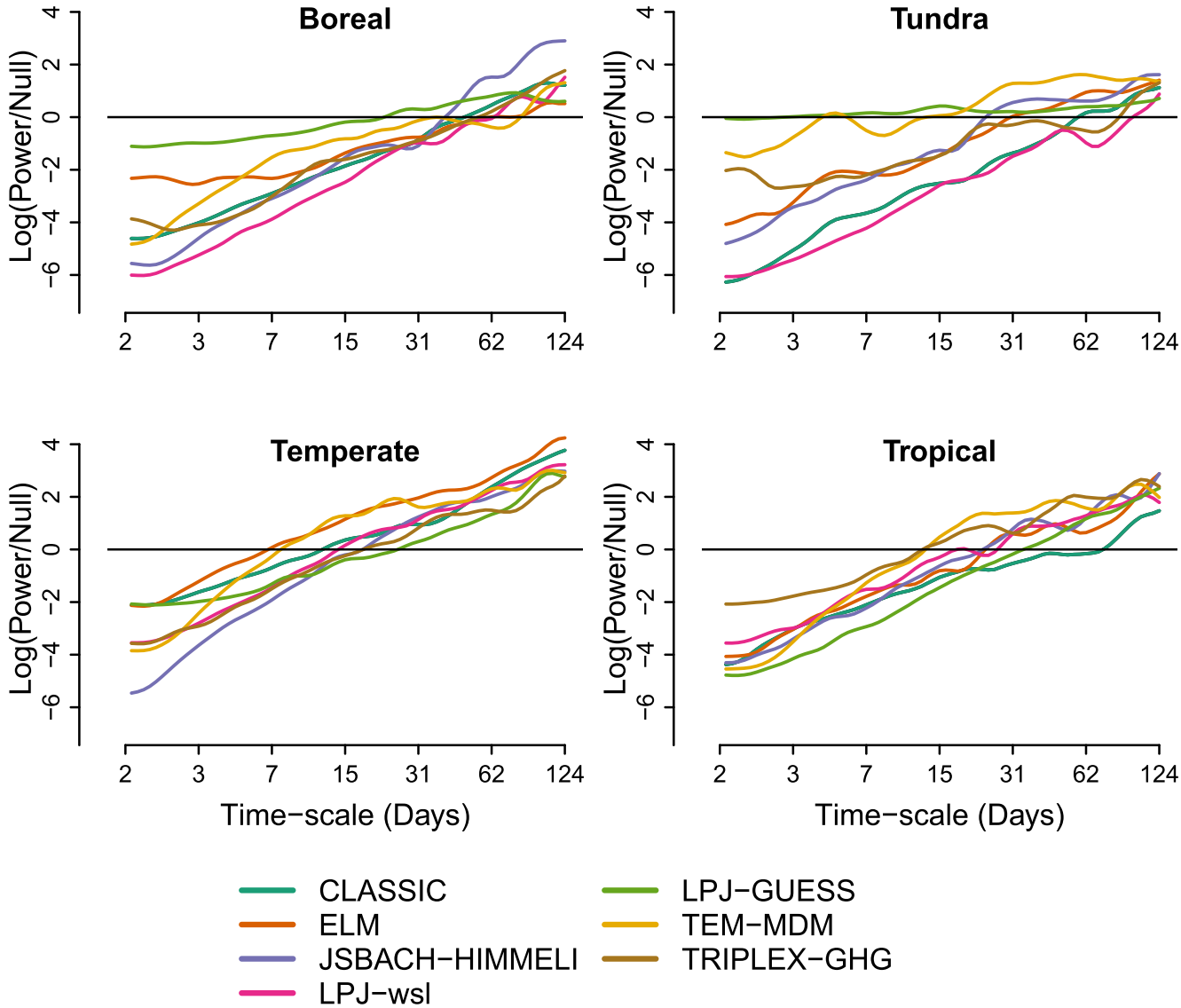

Figure 6. Model performance along time scales with a consideration of uncertainty in observations, reflected by the comparison of model error spectra to the null spectra. The power spectra (Power) are divided by the upper confidence interval of the null (NULL) model (99% quantile) based on logged observations, following the same calculation as Figure 4e. A model error spectrum greater than 0 (horizontal black line) indicates more significant spectral power at these time scales. The slope of the fitted curve represents the scaling component *a*.

frequency (i.e., monthly to seasonal scales). LPJ-wsl tends to underestimate the time span of FCH_4 pulses at a high frequency, with strong pulse emission only occurring in late July, indicating less variability in the modeled FCH_4 during the JJA months. Regarding the disagreement between model and observations, most of the significant regions are in the multi-day to weekly scales, suggesting the model failed to capture the flux variability at these time scales. The discrepancy in FCH_4 only occurs from May to August while it is negligible during the DJF months, when FCH_4 are small and uncertainty is proportionally large.

Figure 4d provides an example of the model-observation mismatch in the global power spectrum for LPJ-wsl and observed FCH_4 at US-WPT in each year separately. Figure 4d is the marginal distribution of the full error spectrum by site-year in Figure 3f, in comparison to the maximum of the spectra of observation error from the Monte Carlo estimates. Here we choose a 99% confidence interval (CI) to define the criteria because, unlike CO_2, FCH_4 is highly spatially heterogeneous and has much higher year-to-year variability. To facilitate the comparison, we divided the model-data error spectra by the 99% CI of the observation error spectra for each time scale

(Figure 4e). Any time scale that falls above the horizontal line (>0) indicates a model residual error that is higher than the uncertainty in the observations. The error of the model is constantly increasing with time-scale, while the random uncertainties in FCH_4 are declining with time-scale with the highest uncertainties at multi-day scales. Here the estimate of the scaling exponent *a* for the LPJ-wsl model at US-WPT sites ranged from 1.5 to 1.7, suggesting a moderate correlation structure (i.e., pink noise).

3.3. Significant Regions of NRE Between the Models and Data

Next, we present the significant regions of model-observation mismatches for all 23 sites and all seven models (Figure 5). Our results suggest that the models have diverse performance with the largest mismatch occurring at the multi-day times scales. For most of the models, the significant mismatch is lower during monthly or seasonal time scales. This pattern confirms the hypothesis that the models generally have better performance in simulating the flux variability at longer time scales than at short-to-intermediate time scales. The increases in significant mismatch at the lowest frequency time scale (>64 days) are likely due to the edge effect, reflecting the limited length of the time series (365 days for a site–year) rather than a confirmation of model performance at capturing fluxes at the time scale. Across the wetland models there are diverse patterns of significant regions in FCH_4, most of which are different from the observation-based patterns, suggesting that there are significant discrepancies between model structures and observed process controls (Figure S1 in Supporting Information S1). The observation-based patterns suggest that most of the significant high power is concentrated within the time scale less than 7 days from May to August (wet season for tropical sites), while the models tend to have relatively high power at a lower frequency (time scale larger than 14 days) at different time periods of the year depending on different model structure.

The comparison of significant regions in model-data mismatch suggests that the models have varying behavior on different wetland types (Figure 5). The majority of models show broadly consistent patterns of significant mismatch across time-scales for the boreal forest and temperate regions. In contrast, the patterns for tropical and subtropical wetland types is diverse among the wetland models. Note that the small sample size of tropical/subtropical wetlands in our study also has an impact on the representativeness of site-level observations. The significant regions for boreal forest and Arctic tundra regions show high power during the JJA seasons (Figure S2 in Supporting Information S1), indicating a consistent dominant control (likely temperature) in the models for these wetlands as suggested by recent studies (Irvin et al., 2021; Knox et al., 2021). For tropical/subtropical wetlands, the significant regions in NRE are spread over all time-scales with diverse patterns across the models, indicating the causes of mismatch with models differ as daily mean temperature becomes less dominant in controlling FCH_4 variability and other processes (e.g., water table dynamics and solar irradiance cycle) become more important.

3.4. Global Model Spectra

We explored the model error patterns by calculating the scaling factor *a* for each model. When considering the observation error in the flux data (the null model is calculated the same way as in Figure 4e), the spectral analysis of the NRE suggests the model errors approximate pink noise patterns for all the wetland models, with the mean scaling exponent *a* of the model estimates ranging from 1.1 to 1.6 for different wetland types (Figure 6). The mean scaling exponent for the boreal forest and Arctic tundra regions (1.1–1.3) was generally lower than that for temperate and tropical regions (1.5–1.6), suggesting the wetland model performance for the temperate and tropical/subtropical regions generally has a longer memory effect (i.e., high tendency for greater persistence of model error) than wetlands in high latitudes. All the models show an increase in error at the monthly and seasonal time scales and the greatest variability across models at multi-day time scales. There was a tendency for the spectral error of some models to exhibit greater persistence than other models. For example, even though the LPJ-wsl model shows relatively low error compared to the other models for boreal and temperate wetlands, the scaling exponents *a* of LPJ-wsl (1.8 and 1.6 respectively) are higher than most of the other models, suggesting that LPJ-wsl model error tends to have a larger memory effect. For the temperate and tropical wetlands, all the models show similar scaling exponents *a* regardless of model structure, indicating the similarity of model behaviors in environmental controls for these wetlands.

Table 3
The p Values of Analysis of Variance Analysis for the Impact of Model Structure on the Spectral Power for Different Wetland Types Within Each of the Four-Time Scales

Wetland type	Time scale	Wetland PFT	Component of CH_4 flux	CH_4 production proxy	Incorporation of nitrogen cycles	Fire	Spatial resolution
Boreal forest	Multiday	<0.001	<0.001	<0.001	<0.001	0.043	ns
	Weekly	<0.001	<0.001	<0.001	<0.001	<0.001	ns
	Monthly	<0.001	<0.001	<0.001	<0.001	<0.001	ns
	Seasonal	<0.001	<0.001	ns	ns	<0.001	ns
Arctic tundra	Multiday	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	Weekly	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	Monthly	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	Seasonal	<0.001	<0.001	<0.001	ns	0.002	0.08
Temperate	Multiday	<0.001	<0.001	ns	<0.001	<0.001	<0.001
	Weekly	<0.001	<0.001	ns	ns	0.004	0.003
	Monthly	<0.001	0.002	ns	0.018	0.003	0.002
	Seasonal	ns	ns	0.001	0.007	<0.001	<0.001
Tropical/ subtropical	Multiday	<0.001	<0.001	<0.001	ns	ns	0.055
	Weekly	<0.001	ns	ns	ns	ns	0.061
	Monthly	<0.001	ns	ns	ns	<0.001	0.019
	Seasonal	<0.001	ns	ns	ns	<0.001	0.002

Note. ns: non significant.

4. Discussion and Conclusions

Our initial hypothesis was that models would perform well at monthly and seasonal time scales because the biogeochemical processes at these time scales are largely driven by solar radiation cycles and corresponding changes in soil temperature. Our results support this hypothesis for Arctic tundra and boreal wetland types where the variations of temperature are the dominant control of FCH_4 (Knox et al., 2021). However, in contrast to our expectations, the models have difficulty capturing variability at monthly and seasonal time scales for temperate and tropical wetlands, where other environmental controls emerge. Considering that the precipitation-driven variables such as water table depth are significantly correlated with the seasonal cycle of FCH_4 at the site level for temperate and tropical sites (Knox et al., 2021), the lower agreements between model and data may be partly caused by discrepancies in precipitation between gridded climate data sets and site-level meteorological conditions. The models also lack representation of hydrological processes at a scale fine enough to reflect the lateral flow from uplands to lowlands and its influence on the water dynamics. The distribution of model wavelet spectra (Figure S1 in Supporting Information S1) on visual inspection appears very different from the site-level measurements, indicating that the models' structures need to better capture variability at multi-day and weekly time scales.

Our analysis further reveals important characteristics in the time series of model errors, which indicates that the errors at short time scales have a memory effect on biases at long time scales. These results suggest that further model development should focus first on correctly replicating flux variability and magnitude at multi-day time scales. Investigations into modeled FCH_4 spectra (Figure S1 in Supporting Information S1) suggest that, in general, models tend to exhibit higher variabilities over monthly time scales from May to August whereas measurements suggest higher variabilities at multi-day time scales. One reason is likely that other environmental variables (e.g., vapor pressure deficit, atmospheric pressure) that regulate FCH_4 variability at short time scales (Stoy et al., 2005) are not included in the model inputs. Additionally, many of the models predict a strong pulse in spectra power across different time scales during a short time period, especially for the JJA months, which causes significant errors at monthly and seasonal time scales (Figures S1 and S2 in Supporting Information S1). This pattern has not been observed by the EC measurements, indicating shared model errors due to the meteorological forcing among models and/or due to missing processes arising from limited understanding of wetland ecosystem dynamics (Neumann et al., 2019; Zona et al., 2016).

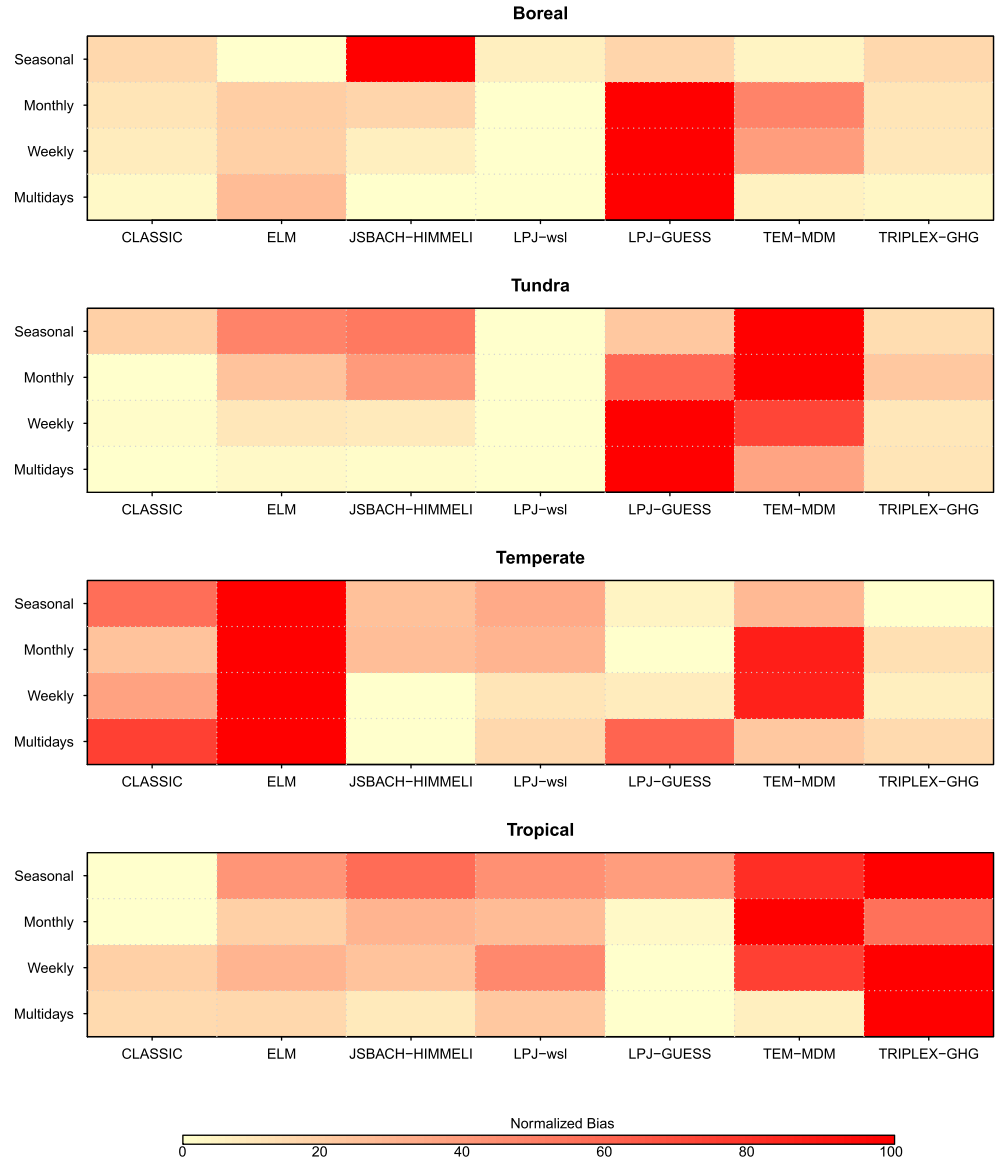

Figure 7. Heat map showing model-observation disagreement by time scales for different wetland types. All of the model-observation disagreement per time scale are normalized to 1–100 with the value of highest model-observation disagreement equal to 100 and lowest to 1. Light yellow and red represent the lowest and highest errors, respectively. The time scales are defined as "Multiday" (2–5 days), "Weekly" (5–15 days), "Monthly" (15–42 days), and "Seasonal scale" (>42 days).

The spectral properties of the model errors along with time scales (Figure 6) indicates that the model structure has an impact on FCH_4 variability, and different groups of models that share similar structure tend to have lower errors propagated from short time scales to high time scales. The ANOVA analysis (Table 3) suggests that the explicit representation of wetland PFTs, CH_4 component fluxes, and wetland production proxies are significantly associated with variance for boreal and Arctic tundra wetland FCH_4 prediction, with a modest and inconsistent effect for temperate and tropical wetlands. The effects of including the nitrogen cycle, fire, and spatial resolution of grid cells were non-significant for most of the time scales. In addition, CH_4 transport through aerenchyma and stomata, which is linked to photosynthesis, and other processes such as ventilation in aerenchymatous vegetation with influence from latent heat flux are critical for models to capture the variability at the diel scale (Knox et al., 2021). Unfortunately, we did not have sub-daily FCH_4 model predictions nor were they driven by site-level meteorological forcings, so we could not evaluate whether representation of processes at the diel scale has an impact on model performance at intermediate scales.

The ranking of model performance across different time scales suggests that no model outperforms others at all time scales and across all wetland types (Figure 7). However, specific models demonstrate better alignment with observed variability across distinct wetland types and time scales, thereby enabling their targeted application to particular regions. Moreover, this quantitative evaluation of model performance offers possibility to improve wetland model ensemble estimations in future studies (Poulter et al., 2017). Given different biogeochemical structures and parameterizations, the analysis suggests inclusion of representation of some key processes in wetland models and proper parameterizations are the basis for improving model performance. However, complex model structure does not guarantee superior model performance, which highlights the importance of properly parameterizing processes at a certain time scale. For instance, models with explicit CH_4 components and multiple wetland PFTs could perform worse than simple models at certain time scales, which is likely due to increased uncertainty from parameterization due to poor observational constraints. A further diagnosis of what environmental and biotic parameters impact the agreement with EC measurements is needed for a better choice of parameter values in representing the realistic temporal variability of FCH_4.

There are a few limitations in the observations affecting our model evaluation. First, the length of observed time series is limited across sites with few sites having more than 5-year records. Unlike CO_2, measurements of FCH_4 are only beginning to cover multiple-year records and thus the EC tower records are not long enough to assess the model's performance in capturing annual and interannual variability. For spectral methods, the short records are particularly problematic for longer sub-annual time scales (e.g., seasonal) due to edge effects on the amount of useable data. Consequently, the magnitude of model-observation disagreement at annual time scales beyond the scope of this study remains uncertain. Given that the wetland model results at annual and interannual time scales are particularly of interest to the global CH_4 budget, having decadal records of measurements is important for an evaluation of model performance at these time scales. Second, the model-site comparisons are statistically challenging as the model-site-year combinations are not randomly distributed but rather depend on the performance at a few sites given the reality of unevenly distributed EC wetland sites. Both undoubtedly have the potential to introduce biases in statistical interpretation and thus influence model score. For instance, the evaluation of model performance for temperate wetlands is strongly affected by model simulations at one marsh site US-TW1 in the United States, which is a restored wetland that contributes ~28% (n = 7 site-years) of the total site-years for temperate wetlands. US-TW1 has a water table height managed at ~25 cm above the soil (Oikawa et al., 2017), which influences the temporal pattern of FCH_4 via hydrological control and thus model evaluations. Lastly, it's worth noting that the scarcity of tropical sites (only 2 sites spanning 5 site-years) has the potential to introduce biases when assessing disagreements between the model and observations. All of the limitations indicate a critical need for more detailed evaluation of model performance at site-level and long-term measurements for underrepresented regions.

One of the important aspects of this analysis is that it is possible that the model performance was underestimated due to the limitation in estimating observation uncertainty and due to potential spatial mismatch between models and EC observations. Although we calculated the spectral uncertainty with the inclusion of observational errors in the evaluation across time scales, the interpretation of whether model-observation disagreements falls outside the acceptable range is strongly influenced by the uncertainty of FCH_4 observations. The default gap-filling methods such as ANN-based estimates for observational uncertainty appear to be overly tight across all sites as suggested by a recent study (Irvin et al., 2021), indicating that actual observation error may be higher than the estimates in our study. In addition, on top of the uncertainty of all the measurements, there is uncertainty originating from a mismatch between the footprints of the individual towers that are usually <1 km^2 and the size of gridded pixels that are often 0.5° or larger (Chu et al., 2021). This footprint mismatch introduces additional noise which is not considered here.

Our study evaluated seven global-scale wetland models from the GCP Methane Assessment against EC FCH_4 measurements from the FLUXNET-CH_4 data set in the time-frequency domain. This analysis helped to identify model-observation disagreements in variability across different time scales and provided guidelines for further wetland model developments. Further detailed intercomparison of model structure and parameterizations is needed to diagnose model structural and parameterization errors. In particular, a more advanced intercomparison protocol would help distinguish structural and parameterization limitations by (a) testing multiple parameterization schemes for major wetland processes (e.g., CH_4 production rate and transport); (b) running the models with inputs from FLUXNET-CH_4 local meteorological condition and local site information such as slope, drainage, and vegetation characteristics; and (c) including longer-term records and spatially representative observations

with full uncertainty characterization from EC tower measurements. In addition, incorporating wavelet analysis into a more comprehensive framework that includes evaluation of other key variables and machine learning-based estimates (Bansal et al., 2023; McNicol et al., 2023) may help identify the factors influencing its performance at specific time scales more effectively. Modeling global-scale wetland CH_4 emissions is essential for accurately quantifying the contribution of wetland-CH_4 feedback to ongoing climate change within the contemporary global CH_4 budget, given their increasing role as potential contributors to the rise in atmospheric CH_4 concentration in recent years (Peng et al., 2022; Zhang et al., 2023). Future intercomparison of wetland CH_4 models will improve understanding of how wetland emissions contribute to variations of atmospheric CH_4 concentration during the past decades and future projections.

Data Availability Statement

The observational data that support the findings of this study are available in the FLUXNET-CH_4 Community Product, available at https://fluxnet.org/data/fluxnet-ch4-community-product/. The modeled results are available at https://doi.org/10.5281/zenodo.7246403.

Acknowledgments

We acknowledge primary support from the Gordon and Betty Moore Foundation (Grant GBMF5439) and from the John Wesley Powell Center for Analysis and Synthesis of the U.S. Geological Survey ("Wetland FLUXNET Synthesis for Methane"). W.J.R., Q.Z, and K.Y.C were supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Scientific Focus Area which is sponsored by the Earth and Environmental Systems Modeling (EESM) Program under the Office of Biological and Environmental Research of the U.S. Department of Energy Office of Science to Lawrence Berkeley National Laboratory (LBNL) under Contract DEAC02-05CH11231. O.S. acknowledges funding by the Canada Research Chairs, Canada Foundation for Innovation Leaders Opportunity Fund, and Natural Sciences and Engineering Research Council Discovery Grant Programs. SB was funded by the U.S. Geological Survey, Ecosystem Mission Area, Land Change Science Climate R&D Program and U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Grant DE-SC0023084. W.Z. acknowledged the support by the Swedish Research Council (Vetenskapsrådet) start Grant (2020-05338). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Arora, V. K., Melton, J. R., & Plummer, D. (2018). An assessment of natural methane fluxes simulated by the CLASS-CTEM model. *Biogeosciences*, *15*(15), 4683–4709. https://doi.org/10.5194/bg-15-4683-2018

Baldocchi, D. (2014). Measuring fluxes of trace gases and energy between ecosystems and the atmosphere – The state and future of the eddy covariance method. *Global Change Biology*, *20*(12), 3600–3609. https://doi.org/10.1111/gcb.12649

Bansal, S., Post van der Burg, M., Fern, R. R., Jones, J. W., Lo, R., McKenna, O. P., et al. (2023). Large increases in methane emissions expected from North America's largest wetland complex. *Science Advances*, *9*(9), eade1112. https://doi.org/10.1126/sciadv.ade1112

Bohrer, G., & Morin, T. H. (2020). *FLUXNET-CH_4 US-ORv Olentangy river wetland Research Park*. The Ohio State University. https://doi.org/10.18140/FLX/1669689

Bousquet, P., Ciais, P., Miller, J. B., Dlugokencky, E. J., Hauglustaine, D. A., Prigent, C., et al. (2006). Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature*, *443*(7110), 439–443. https://doi.org/10.1038/nature05132

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cunha, L. C. D., Cox, P. M., Eliseev, A. V., et al. (2021). Global carbon and other biogeochemical cycles and feedbacks. In *IPCC AR6 WGI, final government distribution (p. Chapter 5)*. Retrieved from https://hal.archives-ouvertes.fr/hal-03336145

Chadburn, S. E., Aalto, T., Aurela, M., Baldocchi, D., Biasi, C., Boike, J., et al. (2020). Modeled microbial dynamics explain the apparent temperature sensitivity of wetland methane emissions. *Global Biogeochemical Cycles*, *34*(11), e2020GB006678. https://doi.org/10.1029/2020GB006678

Chang, K.-Y., Riley, W. J., Knox, S. H., Jackson, R. B., McNicol, G., Poulter, B., et al. (2021). Substantial hysteresis in emergent temperature sensitivity of global wetland CH_4 emissions. *Nature Communications*, *12*(1), 2266. https://doi.org/10.1038/s41467-021-22452-1

Chen, J., & Chu, H. (2020). *FLUXNET-CH_4 US-WPT winous point north marsh*. FluxNet; University of Toledo/Michigan State University. https://doi.org/10.18140/FLX/1669702

Chu, H., Gottgens, J. F., Chen, J., Sun, G., Desai, A. R., Ouyang, Z., et al. (2015). Climatic variability, hydrologic anomaly, and methane emission can turn productive freshwater marshes into net carbon sources. *Global Change Biology*, *21*(3), 1165–1181. https://doi.org/10.1111/gcb.12760

Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., et al. (2021). Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites. *Agricultural and Forest Meteorology*, *301–302*, 108350. https://doi.org/10.1016/j.agrformet.2021.108350

Cornish, C. R., Bretherton, C. S., & Percival, D. B. (2006). Maximal overlap wavelet statistical analysis With application to atmospheric turbulence. *Boundary-Layer Meteorology*, *119*(2), 339–374. https://doi.org/10.1007/s10546-005-9011-y

Dalmagro, H. J., Zanella de Arruda, P. H., Vourlitis, G. L., Lathuillière, M. J., Nogueira, J. D. S., Couto, E. G., & Johnson, M. S. (2019). Radiative forcing of methane fluxes offsets net carbon dioxide uptake for a tropical flooded forest. *Global Change Biology*, *25*(6), 1967–1981. https://doi.org/10.1111/gcb.14615

Delwiche, K. B., Knox, S. H., Malhotra, A., Fluet-Chouinard, E., McNicol, G., Feron, S., et al. (2021). FLUXNET-CH_4: A global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands. *Earth System Science Data*, *13*(7), 3607–3689. https://doi.org/10.5194/essd-13-3607-2021

Desai, A. R., & Thom, J. (2020). *FLUXNET-CH_4 US-Los Lost Creek*. FluxNet; Univ. of Wisconsin. https://doi.org/10.18140/FLX/1669682

Dietze, M. C., Vargas, R., Richardson, A. D., Stoy, P. C., Barr, A. G., Anderson, R. S., et al. (2011). Characterizing the performance of ecosystem models across time scales: A spectral analysis of the North American carbon Program site-level synthesis. *Journal of Geophysical Research*, *116*(G4), G04029s. https://doi.org/10.1029/2011JG001661

Euskirchen, E. (2022a). *AmeriFlux FLUXNET-1F US-BZS Bonanza Creek black spruce*. Lawrence Berkeley National Lab (LBNL), AmeriFlux; University of Alaska Fairbanks, Institute of Arctic Biology. https://doi.org/10.17190/AMF/1881570

Euskirchen, E. (2022b). *AmeriFlux FLUXNET-1F US-BZS Bonanza Creek black spruce*. Lawrence Berkeley National Lab (LBNL), AmeriFlux; University of Alaska Fairbanks, Institute of Arctic Biology. https://doi.org/10.17190/AMF/1881572

Euskirchen, E., Bret-Harte, M., & Edgar, C. (2020). *FLUXNET-CH_4 US-ICs Imnavait Creek watershed wet sedge tundra*. FluxNet; Marine Biological Laboratory; Univ. of Alaska. https://doi.org/10.18140/FLX/1669678

Euskirchen, E., & Edgar, C. (2020). *FLUXNET-CH_4 US-BZB Bonanza Creek thermokarst bog*. FluxNet. https://doi.org/10.18140/FLX/1669668

Friborg, T., & Shurpali, N. (2020). *FLUXNET-CH_4 RU-Vrk Seida/Vorkuta*. FluxNet; Department of Geosciences and Natural Resource Management, University of Copenhagen. https://doi.org/10.18140/FLX/1669658

Golub, M., Koupaei-Abyazani, N., Vesala, T., Mammarella, I., Ojala, A., Bohrer, G., et al. (2023). Diel, seasonal, and inter-annual variation in carbon dioxide effluxes from lakes and reservoirs. *Environmental Research Letters*, *18*(3), 034046. https://doi.org/10.1088/1748-9326/acb834

Gouhier, T. C., Grinsted, A., & Simko, V. (2021). R package biwavelet: Conduct Univariate and Bivariate wavelet analyses. Retrieved from https://github.com/tgouhier/biwavelet

Hollinger, D. Y., & Richardson, A. D. (2005). Uncertainty in eddy covariance measurements and its application to physiological models. *Tree Physiology*, *25*(7), 873–885. https://doi.org/10.1093/treephys/25.7.873

IPCC. (2013). *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental panel on climate change*. Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar5/wg1/

Irvin, J., Zhou, S., McNicol, G., Lu, F., Liu, V., Fluet-Chouinard, E., et al. (2021). Gap-filling eddy covariance methane fluxes: Comparison of machine learning model predictions and uncertainties at FLUXNET-CH_4 wetlands. *Agricultural and Forest Meteorology*, *308–309*, 108528. https://doi.org/10.1016/j.agrformet.2021.108528

Iwata, H., Ueyama, M., & Harazono, Y. (2020). *FLUXNET-CH_4 US-Uaf University of Alaska, Fairbanks*. FluxNet; Osaka Prefecture University; Shinshu University. https://doi.org/10.18140/FLX/1669701

Jansen, J., Friborg, T., Jammet, M., & Crill, P. (2020). *FLUXNET-CH_4 SE-St1 Stordalen grassland*. FluxNet; University of Copenhagen. https://doi.org/10.18140/FLX/1669660

Knox, S. H., Bansal, S., McNicol, G., Schafer, K., Sturtevant, C., Ueyama, M., et al. (2021). Identifying dominant environmental predictors of freshwater wetland methane fluxes across diurnal to seasonal time scales. *Global Change Biology*, *27*(15), 3582–3604. https://doi.org/10.1111/gcb.15661

Knox, S. H., Jackson, R. B., Poulter, B., McNicol, G., Fluet-Chouinard, E., Zhang, Z., et al. (2019). FLUXNET-CH_4 synthesis activity: Objectives, observations, and future directions. *Bulletin of the American Meteorological Society*, *100*(12), 2607–2632. https://doi.org/10.1175/BAMS-D-18-0268.1

Lasslop, G., Reichstein, M., Kattge, J., & Papale, D. (2008). Influences of observation errors in eddy flux data on inverse model parameter estimation. *Biogeosciences*, *5*(5), 1311–1324. https://doi.org/10.5194/bg-5-1311-2008

Li, Y., Wu, Y., Tang, J., Zhu, P., Gao, Z., & Yang, Y. (2023). Quantitative evaluation of wavelet analysis method for turbulent flux calculation of non-stationary series. *Geophysical Research Letters*, *50*(5), e2022GL101591. https://doi.org/10.1029/2022GL101591

Liu, L., Zhuang, Q., Oh, Y., Shurpali, N. J., Kim, S., & Poulter, B. (2020). Uncertainty quantification of global net methane emissions from terrestrial ecosystems using a mechanistically based biogeochemistry model. *Journal of Geophysical Research: Biogeosciences*, *125*(6), e2019JG005428. https://doi.org/10.1029/2019JG005428

Liu, Y., Liang, X. S., & Weisberg, R. H. (2007). Rectification of the bias in the wavelet power spectrum. *Journal of Atmospheric and Oceanic Technology*, *24*(12), 2093–2102. https://doi.org/10.1175/2007JTECHO511.1

Lohila, A., Aurela, M., Tuovinen, J.-P., Laurila, T., Hatakka, J., Rainne, J., & Mäkelä, T. (2020). FLUXNET-CH_4 FI-Lom Lompolojankka [Dataset]. https://doi.org/10.18140/FLX/1669638

Lunt, M. F., Palmer, P. I., Feng, L., Taylor, C. M., Boesch, H., & Parker, R. J. (2019). An increase in methane emissions from tropical Africa between 2010 and 2016 inferred from satellite data. *Atmospheric Chemistry and Physics*, *19*(23), 14721–14740. https://doi.org/10.5194/acp-19-14721-2019

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J., et al. (2021). 2010–2015 North American methane emissions, sectoral contributions, and trends: A high-resolution inversion of GOSAT observations of atmospheric methane. *Atmospheric Chemistry and Physics*, *21*(6), 4339–4356. https://doi.org/10.5194/acp-21-4339-2021

McGuire, A. D., Christensen, T. R., Hayes, D., Heroult, A., Euskirchen, E., Kimball, J. S., et al. (2012). An assessment of the carbon balance of Arctic tundra: Comparisons among observations, process models, and atmospheric inversions. *Biogeosciences*, *9*(8), 3185–3204. https://doi.org/10.5194/bg-9-3185-2012

McNicol, G., Fluet-Chouinard, E., Ouyang, Z., Knox, S., Zhang, Z., Aalto, T., et al. (2023). Upscaling wetland methane emissions from the FLUXNET-CH4 eddy covariance network (UpCH4 v1.0): Model development, network Assessment, and Budget Comparison. *AGU Advances*, *4*(5), e2023AV000956. https://doi.org/10.1029/2023AV000956

Melton, J. R., & Arora, V. K. (2016). Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (CTEM) v. 2.0. *Geoscientific Model Development*, *9*(1), 323–361. https://doi.org/10.5194/gmd-9-323-2016

Melton, J. R., Wania, R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., et al. (2013). Present state of global wetland extent and wetland methane modelling: Conclusions from a model inter-comparison project (WETCHIMP). *Biogeosciences*, *10*(2), 753–788. https://doi.org/10.5194/bg-10-753-2013

Meyers, S. D., Kelly, B. G., & O'Brien, J. J. (1993). An introduction to wavelet analysis in oceanography and meteorology: With application to the dispersion of Yanai waves. *Monthly Weather Review*, *121*(10), 2858–2866. https://doi.org/10.1175/1520-0493(1993)121<2858:AITWAI>2.0.CO;2

Neumann, R. B., Moorberg, C. J., Lundquist, J. D., Turner, J. C., Waldrop, M. P., McFarland, J. W., et al. (2019). Warming effects of spring rainfall increase methane emissions from thawing permafrost. *Geophysical Research Letters*, *46*(3), 1393–1401. https://doi.org/10.1029/2018gl081274

Nilsson, M. B., & Peichl, M. (2020). *FLUXNET-CH_4 SE-Deg Degero*. FluxNet; Department of Forest Ecology and Management; Swedish University of Agricultural Sciences. https://doi.org/10.18140/FLX/1669659

Oikawa, P. Y., Jenerette, G. D., Knox, S. H., Sturtevant, C., Verfaillie, J., Dronova, I., et al. (2017). Evaluation of a hierarchy of models reveals importance of substrate limitation for predicting carbon dioxide and methane exchange in restored wetlands. *Journal of Geophysical Research: Biogeosciences*, *122*(1), 145–167. https://doi.org/10.1002/2016JG003438

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial ecoregions of the world: A new map of life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51*(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Pallandt, M., Kumar, J., Mauritz, M., Schuur, E., Virkkala, A.-M., Celis, G., et al. (2021). Representativeness assessment of the pan-Arctic eddy-covariance site network, and optimized future enhancements. *Biogeosciences Discussions*, 1–42. https://doi.org/10.5194/bg-2021-133

Peltola, O., Hensen, A., Belelli Marchesini, L., Helfter, C., Bosveld, F. C., van den Bulk, W. C. M., et al. (2015). Studying the spatial variability of methane flux with five eddy covariance towers of varying height. *Agricultural and Forest Meteorology*, *214–215*, 456–472. https://doi.org/10.1016/j.agrformet.2015.09.007

Peng, S., Lin, X., Thompson, R. L., Xi, Y., Liu, G., Hauglustaine, D., et al. (2022). Wetland emission and atmospheric sink changes explain methane growth in 2020. *Nature*, *612*(7940), 477–482. https://doi.org/10.1038/s41586-022-05447-w

Poulter, B., Bousquet, P., Canadell, J. G., Ciais, P., Peregon, A., Saunois, M., et al. (2017). Global wetland contribution to 2000–2012 atmospheric methane growth rate dynamics. *Environmental Research Letters*, *12*(9), 094013. https://doi.org/10.1088/1748-9326/aa8391

Raivonen, M., Smolander, S., Backman, L., Susiluoto, J., Aalto, T., Markkanen, T., et al. (2017). HIMMELI v1.0: HelsinkI Model of MEthane buiLd-up and emIssionfor peatlands. *Geoscientific Model Development Discussions*, 1–45. https://doi.org/10.5194/gmd-2017-52

Richardson, A. D., Anderson, R. S., Arain, M. A., Barr, A. G., Bohrer, G., Chen, G., et al. (2012). Terrestrial biosphere models need better representation of vegetation phenology: Results from the North American carbon Program site synthesis. *Global Change Biology*, *18*(2), 566–584. https://doi.org/10.1111/j.1365-2486.2011.02562.x

Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K. J., Flanagan, L. B., Katul, G. G., et al. (2006). A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. *Agricultural and Forest Meteorology*, *136*(1), 1–18. https://doi.org/10.1016/j.agrformet.2006.01.007

Richardson, A. D., Mahecha, M. D., Falge, E., Kattge, J., Moffat, A. M., Papale, D., et al. (2008). Statistical properties of random CO_2 flux measurement uncertainty inferred from model residuals. *Agricultural and Forest Meteorology*, *148*(1), 38–50. https://doi.org/10.1016/j.agrformet.2007.09.001

Riley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., et al. (2011). Barriers to predicting changes in global terrestrial methane fluxes: Analyses using CLM4Me, a methane biogeochemistry model integrated in CESM. *Biogeosciences*, *8*(7), 1925–1953. https://doi.org/10.5194/bg-8-1925-2011

Ringeval, B., Houweling, S., van Bodegom, P. M., Spahni, R., van Beek, R., Joos, F., & Röckmann, T. (2014). Methane emissions from floodplains in the Amazon basin: Challenges in developing a process-based model for global applications. *Biogeosciences*, *11*(6), 1519–1558. https://doi.org/10.5194/bg-11-1519-2014

Saunois, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J. G., et al. (2017). Variability and quasi-decadal changes in the methane budget over the period 2000–2012. *Atmospheric Chemistry and Physics*, *17*(18), 11135–11161. https://doi.org/10.5194/acp-17-11135-2017

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., et al. (2020). The global methane budget 2000–2017. *Earth System Science Data*, *12*(3), 1561–1623. https://doi.org/10.5194/essd-12-1561-2020

Schaefer, K., Schwalm, C. R., Williams, C., Arain, M. A., Barr, A., Chen, J. M., et al. (2012). A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research*, *117*(G3), G03010. https://doi.org/10.1029/2012JG001960

Schuur, E. A. (2020). *FLUXNET-CH_4 US-EML eight mile lake permafrost thaw gradient, Healy Alaska*. FluxNet; Northern Arizona Univ; University of Northern Arizona. https://doi.org/10.18140/FLX/1669674

Schwalm, C. R., Williams, C. A., Schaefer, K., Anderson, R., Arain, M. A., Baker, I., et al. (2010). A model-data intercomparison of CO_2 exchange across North America: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research*, *115*(G3), G00H05. https://doi.org/10.1029/2009JG001229

Shortt, R., Hemes, K., Szutu, D., Verfaillie, J., & Baldocchi, D. (2020). *FLUXNET-CH_4 US-Sne Sherman Island restored wetland*. FluxNet; Univ. of California, Berkeley. https://doi.org/10.18140/FLX/1669693

Sonnentag, O., & Helbig, M. (2020). FLUXNET-CH_4 CA-SCB Scotty Creek bog (2014-2017) [Dataset]. https://doi.org/10.18140/FLX/1669613

Stavert, A. R., Saunois, M., Canadell, J. G., Poulter, B., Jackson, R. B., Regnier, P., et al. (2021). Regional trends and drivers of the global methane budget. *Global Change Biology*, *28*(1), 182–200. https://doi.org/10.1111/gcb.15901

Stoy, P. C., Dietze, M. C., Richardson, A. D., Vargas, R., Barr, A. G., Anderson, R. S., et al. (2013). Evaluating the agreement between measurements and models of net ecosystem exchange at different times and timescales using wavelet coherence: An example using data from the North American carbon program site-level interim synthesis. *Biogeosciences*, *10*(11), 6893–6909. https://doi.org/10.5194/bg-10-6893-2013

Stoy, P. C., Katul, G. G., Siqueira, M. B. S., Juang, J.-Y., McCarthy, H. R., Kim, H.-S., et al. (2005). Variability in net ecosystem exchange from hourly to inter-annual time scales at adjacent pine and hardwood forests: A wavelet analysis. *Tree Physiology*, *25*(7), 887–902. https://doi.org/10.1093/treephys/25.7.887

Sturtevant, C., Ruddell, B. L., Knox, S. H., Verfaillie, J., Matthes, J. H., Oikawa, P. Y., & Baldocchi, D. (2016). Identifying scale-emergent, nonlinear, asynchronous processes of wetland methane exchange. *Journal of Geophysical Research: Biogeosciences*, *121*(1), 188–204. https://doi.org/10.1002/2015JG003054

Tao, J., Zhu, Q., Riley, W. J., & Neumann, R. B. (2021). Improved ELMv1-ECA simulations of zero-curtain periods and cold-season CH_4 and CO_2 emissions at Alaskan Arctic tundra sites. *The Cryosphere*, *15*(12), 5281–5307. https://doi.org/10.5194/tc-15-5281-2021

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Valach, A., Szutu, D., Eichelmann, E., Knox, S., Verfaillie, J., & Baldocchi, D. (2020). *FLUXNET-CH_4 US-tw1 twitchell wetland west pond*. FluxNet; Univ. of California, Berkeley. https://doi.org/10.18140/FLX/1669696

Vargas, R., Detto, M., Baldocchi, D. D., & Allen, M. F. (2010). Multiscale analysis of temporal variability of soil CO_2 production as influenced by weather and vegetation. *Global Change Biology*, *16*(5), 1589–1605. https://doi.org/10.1111/j.1365-2486.2009.02111.x

Wania, R., Melton, J. R., Hodson, E. L., Poulter, B., Ringeval, B., Spahni, R., et al. (2013). Present state of global wetland extent and wetland methane modelling: Methodology of a model inter-comparison project (WETCHIMP). *Geoscientific Model Development*, *6*(3), 617–641. https://doi.org/10.5194/gmd-6-617-2013

Wania, R., Ross, I., & Prentice, I. C. (2009). Integrating peatlands and permafrost into a dynamic global vegetation model: 2. Evaluation and sensitivity of vegetation and carbon cycle processes: Peatlands and permafrost in LPJ, 2. *Global Biogeochemical Cycles*, *23*(3), GB3015. https://doi.org/10.1029/2008GB003413

Wania, R., Ross, I., & Prentice, I. C. (2010). Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3.1. *Geoscientific Model Development*, *3*(2), 565–584. https://doi.org/10.5194/gmd-3-565-2010

Wong, G., Melling, L., Tang, A., Aeries, E., Waili, J., Musin, K., et al. (2020). FLUXNET-CH_4 MY-MLM Maludam National park [Dataset]. https://doi.org/10.18140/FLX/1669650

Xu, X., Riley, W. J., Koven, C. D., Billesbach, D. P., Chang, R. Y.-W., Commane, R., et al. (2016). A multi-scale comparison of modeled and observed seasonal methane emissions in northern wetlands. *Biogeosciences*, *13*(17), 5043–5056. https://doi.org/10.5194/bg-13-5043-2016

Xu, X., Yuan, F., Hanson, P. J., Wullschleger, S. D., Thornton, P. E., Riley, W. J., et al. (2016). Reviews and syntheses: Four decades of modeling methane cycling in terrestrial ecosystems. *Biogeosciences Discussions*, 1–56. https://doi.org/10.5194/bg-2016-37

Yu, X., Millet, D. B., Wells, K. C., Henze, D. K., Cao, H., Griffis, T. J., et al. (2021). Aircraft-based inversions quantify the importance of wetlands and livestock for Upper Midwest methane emissions. *Atmospheric Chemistry and Physics*, *21*(2), 951–971. https://doi.org/10.5194/acp-21-951-2021

Zhang, Z., Fluet-Chouinard, E., Jensen, K., McDonald, K., Hugelius, G., Gumbricht, T., et al. (2021). Development of the global dataset of wetland area and dynamics for methane modeling (WAD2M). *Earth System Science Data*, *13*(5), 2001–2023. https://doi.org/10.5194/essd-13-2001-2021

Zhang, Z., Poulter, B., Feldman, A. F., Ying, Q., Ciais, P., Peng, S., & Li, X. (2023). Recent intensification of wetland methane feedback. *Nature Climate Change*, *13*(5), 430–433. https://doi.org/10.1038/s41558-023-01629-0

Zhang, Z., Poulter, B., Knox, S., Stavert, A., McNicol, G., Fluet-Chouinard, E., et al. (2021). Anthropogenic emission is the main contributor to the rise of atmospheric methane during 1993–2017. *National Science Review*, *9*(5), nwab200. https://doi.org/10.1093/nsr/nwab200

Zhang, Z., Zimmermann, N. E., Calle, L., Hurtt, G., Chatterjee, A., & Poulter, B. (2018). Enhanced response of global wetland methane emissions to the 2015–2016 El Niño-Southern Oscillation event. *Environmental Research Letters*, *13*(7), 074009. https://doi.org/10.1088/1748-9326/aac939

Zhang, Z., Zimmermann, N. E., Kaplan, J. O., & Poulter, B. (2016). Modeling spatiotemporal dynamics of global wetlands: Comprehensive evaluation of a new sub-grid TOPMODEL parameterization and uncertainties. *Biogeosciences*, *13*(5), 1387–1408. https://doi.org/10.5194/bg-13-1387-2016

Zhu, Q., Peng, C., Chen, H., Fang, X., Liu, J., Jiang, H., et al. (2015). Estimating global natural wetland methane emissions using process modelling: Spatio-temporal patterns and contributions to atmospheric methane fluctuations: Global natural wetland methane emissions. *Global Ecology and Biogeography*, *24*(8), 959–972. https://doi.org/10.1111/geb.12307

Zhuang, Q., Chen, M., Xu, K., Tang, J., Saikawa, E., Lu, Y., et al. (2013). Response of global soil consumption of atmospheric methane to changes in atmospheric climate and nitrogen deposition: Global soil consumption of methane. *Global Biogeochemical Cycles*, *27*(3), 650–663. https://doi.org/10.1002/gbc.20057

Zhuang, Q., Melillo, J. M., Kicklighter, D. W., Prinn, R. G., McGuire, A. D., Steudler, P. A., et al. (2004). Methane fluxes between terrestrial ecosystems and the atmosphere at northern high latitudes during the past century: A retrospective analysis with a process-based biogeochemistry model. *Global Biogeochemical Cycles*, *18*(3), GB3010. https://doi.org/10.1029/2004GB002239

Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., et al. (2016). Cold season emissions dominate the Arctic tundra methane budget. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(1), 40–45. https://doi.org/10.1073/pnas.1516017113

Zona, D., & Oechel, W. C. (2020a). *FLUXNET-CH_4 US-Atq Atqasuk*. FluxNet; San Diego State Univ. https://doi.org/10.18140/FLX/1669663

Zona, D., & Oechel, W. C. (2020b). *FLUXNET-CH_4 US-Ivo Ivotuk*. FluxNet; San Diego State Univ. https://doi.org/10.18140/FLX/1669679

Erratum

The originally published version of this article contained errors in the affiliations of two co-authors, Antti Leppänen and Maarit Raivonen. The correct affiliation for both authors is Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, Helsinki, Finland. The errors have been corrected, and this may be considered the authoritative version of record.